

Spektralanalytische Untersuchungsverfahren der Sprache

Hans Goydke

Zusammenfassung Zur Spektralanalyse von Sprache können verschiedene Verfahren und Analysator-Typen eingesetzt werden. Es werden Parallelfilter-, Suchton- und FFT-Analysatoren beschrieben und ihre speziellen Eigenschaften in bezug auf die Analyse von Sprachsignalen an Hand von Beispielen demonstriert.

1. Einleitung

Sprachlaute können sehr unterschiedlichen Signalcharakter aufweisen. Bei einer Fourieranalyse, bei der die Signale in einfache Schwingungsformen zerlegt werden, sind daher stets mehrere Bedingungen zu beachten, um zu einem, jeweils unter einem bestimmten Gesichtspunkt optimalen Analyseergebnis zu kommen.

Für die verschiedenen Anwendungsfälle stehen heute unterschiedliche Analyse-Verfahren und Analysator-Typen zur Verfügung, die zu recht unterschiedlichen Ergebnissen führen können. Die Ursachen hierfür sind häufig vielfältig und schwer erkennbar. Mit Hilfe von Beispielen aus der 1969 erfolgten Neuaufnahme der »Wörter für Gehörprüfung mit Sprache« [1] (nach DIN 45621) werden die Besonderheiten der heute hauptsächlich gebräuchlichen spektralanalytischen Verfahren aufgezeigt.

2. Theorien der Spektralanalyse

Bei einer Spektralanalyse nach Fourier wird die Zeitfunktion $f(t)$ in eine Frequenzfunktion $\mathfrak{F}(\omega)$ umgewandelt. Periodische, nicht sinusförmige Schwingungen werden in sinusförmige Teilschwingungen zerlegt, deren Frequenzen ganzzahlige Vielfache der tiefsten vorkommenden Frequenz, der Grundfrequenz, sind. Die Fourieranalyse legt nicht nur die Amplituden der Teilschwingungen, sondern auch ihre Phasen fest. Dem Phasenspektrum wird bei der Sprachanalyse bisher wenig Beachtung geschenkt, da einmal die Phasenlagen der Teiltöne für den Höreindruck eines Signals im allgemeinen keine Rolle spielen, zum andern ergibt sich die Schwierigkeit, daß die Phasenwinkel in ihrer absoluten Größe und in ihrer Lage zueinander von der Wahl des Zeitpunktes $t=0$ abhängen.

Das Spektrum ist im allgemeinen komplex und ergibt sich zu:

$$\mathfrak{F}(\omega) = \int_{-\infty}^{+\infty} f(t) \cdot e^{-j\omega t} dt \quad (1)$$

und läßt sich umformen in

$$\mathfrak{F}(\omega) = \int_{-\infty}^{+\infty} f(t) \cdot \cos \omega t dt - j \int_{-\infty}^{+\infty} f(t) \cdot \sin \omega t dt \quad (2)$$

Mitteilung aus der Physikalisch-Technischen Bundesanstalt (PTB), Braunschweig

1. Introduction

Speech sounds can vary greatly in their signal character. In a Fourier analysis, in which the signals are analyzed into simple forms of oscillations a number of conditions must always be taken into consideration, in order to obtain in each case an analysis result that is optimal for the particular point of view predominating.

Today, various different methods and types of analyzers are available for the various cases of application, and they can lead to widely differing results. The reasons for this are frequently multiple and difficult to recognize. With the aid of examples from the "Words for Testing Hearing with Speech", newly recorded in 1969 [1] (in accordance with DIN 45621), the peculiarities of the spectral analytical methods that are mainly used today are demonstrated.

2. Theories of spectral analysis

In a spectral analysis of the Fourier type, the time function $f(t)$ is converted into a frequency function $\mathfrak{F}(\omega)$. Periodic, non sinusoidal oscillations are split into sinusoidal partial oscillations, their frequencies are multiples of the lowest frequency obtaining, the basic frequency. The Fourier analysis determines not only the amplitudes of the partial oscillations, but also their phases. To date, in speech analysis little attention has been paid to the phase spectrum as, on the one hand, the phase relations of the partial tones in general play no role for the hearing impression of a signal, and, on the other hand, a difficulty is involved, namely, that the phase angles in their absolute magnitude and their relationship to one another, depend upon the selection of the time point $t = 0$.

In general, the spectrum is complex and can be expressed by:

$$\mathfrak{F}(\omega) = \int_{-\infty}^{+\infty} f(t) \cdot e^{-j\omega t} dt \quad (1)$$

and can be converted to

$$\mathfrak{F}(\omega) = \int_{-\infty}^{+\infty} f(t) \cdot \cos \omega t dt - j \int_{-\infty}^{+\infty} f(t) \cdot \sin \omega t dt \quad (2)$$

Report from the Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Federal Republic of Germany

Methods of Speech Spectral Analysis

Summary For the spectral analysis of speech, various methods and types of analyzers can be employed. Parallel filter, heterodyne- and FFT-analyzers are described and their specific properties with respect to analysis of speech signals are demonstrated with the aid of examples.

Betrachtet man die Zeitfunktion eines Sprachsignals zum Beispiel am Oszillogramm des Wortes »Spott« in Abb. 1, erkennt man bereits den unterschiedlichen Signalcharakter einzelner Sprachlaute. Dem Zischlaut »S« mit rauschartigem Charakter folgt eine kleine Pause, nach der Freigabe des Luftstromverschlusses »p« im Spracherzeugungstrakt erscheint der Vokal »o« mit einer deutlich periodischen Zeitfunktion. Es schließt sich eine Pause mit nachfolgendem Explosivlaut »t« bei der plötzlichen Verschlussfreigabe an, wobei sich wiederum als Zeitfunktion ein Rauschsignal zeigt.

Die Länge des Vokals läßt sich hier mit ca. 100 ms abschätzen. Stimmhafte Laute können in der Länge je nach Sprechweise und Lautumgebung um den Faktor 10 schwanken. Als kürzeste Dauer stimmhafter Laute, bei der eine Erkennbarkeit gesichert ist, wurde für die deutsche Sprache 4 bis 30 ms ermittelt, für stimmlose Laute 60 bis 120 ms [2]. Hinzu kommt, daß bei kontinuierlichen Sprachsignalen die Sprachlaute ineinander übergehen. Es eignet sich damit die Fouriertransformation nach Gleichung (1) nicht ohne weiteres zur Analyse einzelner Sprachlaute, da die Transformation der

If we observe the time function of a speech signal, for example the oscillogram of the word "Spott" in Fig. 1, we can already recognize the differing signal character of individual speech sounds. The fricative "S", with the noise-like character is followed by a small pause; after releasing the air-flow block "p" in the speech-producing tract, the vowel "o" appears with a clearly periodic time function. There then follows a pause with the subsequent plosive "t" on the sudden release of the blockage, so that again the time function in form of a noise signal appears.

The length of the vowel can be estimated here at about 100 ms. Depending upon the mode of speaking and the ambient sound voiced sounds can vary in their length by a factor of 10. For the German language, 4 to 30 ms has been established as the shortest duration of voiced sounds, at which recognizability is still ensured, for unvoiced sounds, 60 to 120 ms [2]. A further point is that in continuous speech signals, the speech sounds "flow into one another". Thus, Fourier transformation in accordance with equation (1) is not necessarily suitable for the analysis of single speech sounds,

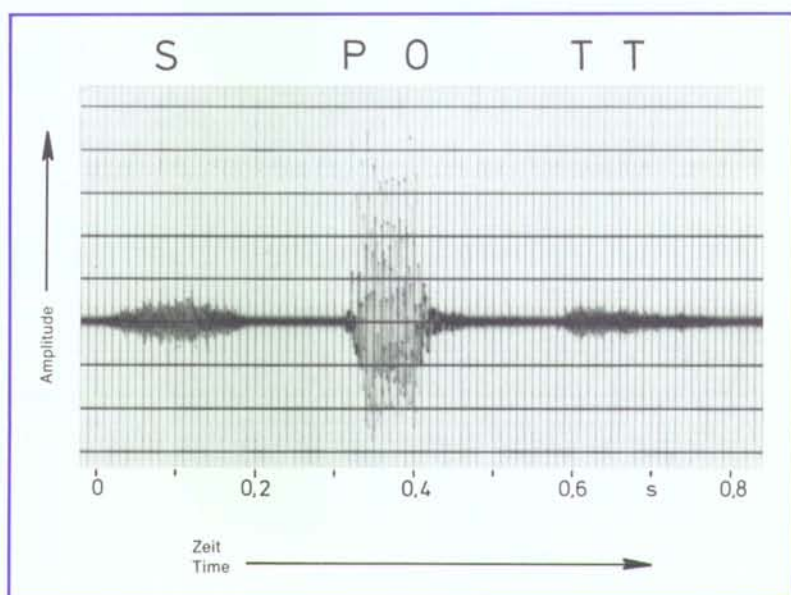


Abb. 1 Zeitfunktion des Wortes »Spott«

Fig. 1 Time function of the word "Spott"

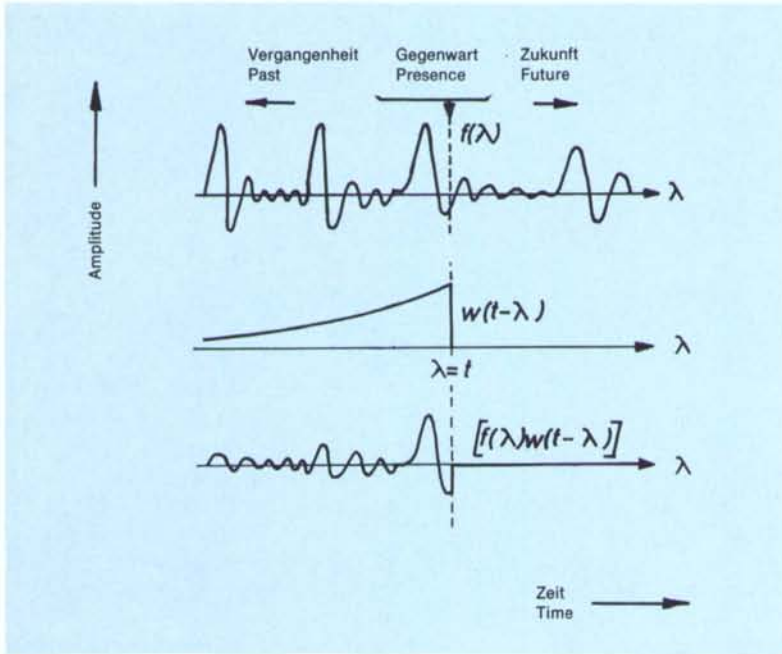


Abb. 2 Bewertung der Zeitfunktion $f(t)$ durch die Bewertungsfunktion $w(t)$

Fig. 2 Weighting of the time function $f(t)$ by means of the weighting function $w(t)$

gesamten Zeitfunktion von $-\infty$ bis $+\infty$ in ein einziges Spektrum vorgenommen wird. Benötigt wird ein fortlaufend zu bildendes Kurzzeitspektrum, das jeweils aus den vergangenen Werten der Zeitfunktion bestimmt werden kann und bei dem die Zeit eine unabhängige Variable bildet. Dies läßt sich erreichen, indem die Zeitfunktion in geeigneter Weise mit einer Bewertungsfunktion $w(t)$ amplitudenbewertet wird, wie es in Abb. 2 veranschaulicht ist. Das Kurzzeitspektrum ergibt sich dann zu

$$\begin{aligned} \mathfrak{F}(\omega, t) &= \int_{-\infty}^t f(\lambda) \cdot \cos \omega \lambda \cdot w(t - \lambda) d\lambda - \\ & j \int_{-\infty}^t f(\lambda) \cdot \sin \omega \lambda \cdot w(t - \lambda) d\lambda \\ &= a(\omega, t) - j b(\omega, t). \end{aligned} \quad (3)$$

Für die Untersuchung von Sprache beschränkt man sich im allgemeinen auf die Bestimmung des Betrages

$$\begin{aligned} F(\omega, t) &= |\mathfrak{F}(\omega, t)| \\ &= \sqrt{a^2 + b^2} \end{aligned} \quad (4)$$

Bei der technischen Realisierung dieser Formeln findet das in Abb. 3 gezeigte einfache Prinzip praktische Anwendung. Es

since the transformation of the total time function from $-\infty$ to $+\infty$ is carried out in a single spectrum. What is needed is a continuously created short-time spectrum which can be determined in each case from the past values of the time function, and in which the time represents an independent variable. This can be achieved by amplitude-weighting the time function in a suitable manner employing a weighting function $w(t)$, as is shown in Fig. 2. The short-time spectrum is then expressed as

$$\begin{aligned} \mathfrak{F}(\omega, t) &= \int_{-\infty}^t f(\lambda) \cdot \cos \omega \lambda \cdot w(t - \lambda) d\lambda - \\ & j \int_{-\infty}^t f(\lambda) \cdot \sin \omega \lambda \cdot w(t - \lambda) d\lambda \\ &= a(\omega, t) - j b(\omega, t). \end{aligned} \quad (3)$$

For the examination of speech, the limitation is usual to the determination of the quantity

$$\begin{aligned} F(\omega, t) &= |\mathfrak{F}(\omega, t)| \\ &= \sqrt{a^2 + b^2}. \end{aligned} \quad (4)$$

For the technical realization of this formula the simple principle shown in Fig. 3 finds practical application. It can be

Abb. 3 Schaltungsprinzip von Filteranalysatoren

Fig. 3 Circuitry principle of filter analyzers

Bandpaßfilter = bandpass filter
Gleichrichter = rectifier
Tiefpaßfilter = low-pass filter

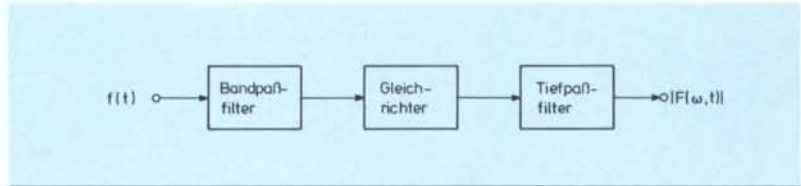
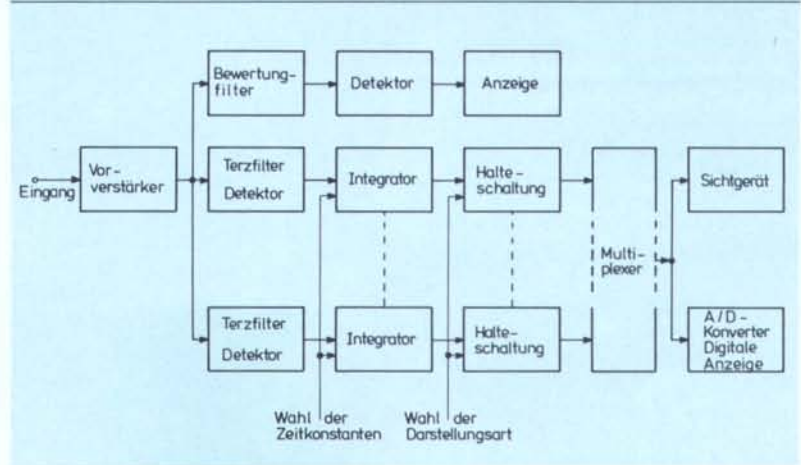


Abb. 4 Schaltungsprinzip eines Echtzeit-Terzfilter-Analysators

Fig. 4 Circuitry principle of a real-time 1/3 octave filter analyzer

Eingang = input
Vorverstärker = preamplifier
Bewertungsfiler = weighting filter
Detektor = detector
Anzeige = indicator
Terzfilter = 1/3 octave filter
Halteschaltung = Hold circuit
Wahl der Zeitkonstanten = time constant mode
Wahl der Darstellungsart = display mode
Sichtgerät = picture screen
A/D-Konverter = A/D-converter
Digitale Anzeige = digital display



läßt sich nämlich mathematisch nachweisen [3], daß bereits die Umhüllende der Ausgangsspannung eines Bandfilters, an dessen Eingang die Zeitfunktion $f(t)$ liegt, den zeitlichen Verlauf des Betrages des Kurzzeitspektrums für die Frequenz liefert, auf die das Filter abgestimmt ist. Die Umhüllende wird gewonnen durch einen dem Filter nachgeschalteten Gleichrichter, dem ein Glättungstiefpaß folgt. Als Bewertungsfunktion wird die Umhüllende der Impulsantwort des Filters wirksam. Im Spektrum erscheint damit nicht nur die Fouriertransformierte der Zeitfunktion, sondern zusätzlich die der Bewertungsfunktion. Bei der Fouriertransformation nach Gleichung (1) einer reinen Sinusschwingung würde sich im Spektrum eine Linie zeigen. Beim Kurzzeitspektrum, gebildet mit einem auf die Frequenz der Sinusschwingung abgestimmten Bandfilter, »entartet« die Linie zur »Filterdurchlaßkurve« im Spektrum. Die Fouriertransformierte eines rechteckförmigen Signals ist die Spaltfunktion $\frac{\sin x}{x}$.

Beim idealen Bandfilter mit rechteckförmiger Filterdurchlaßkurve ist also die Bewertungsfunktion $w(t)$ im Zeitbereich gleich der Spaltfunktion. Entsprechend zeigt sich bei der mathematischen Fouriertransformation eines rechteckigen Ausschnittes aus einem Sinussignal an Stelle der Spektrallinie die Spaltfunktion.

demonstrated mathematically [3] that already the envelope of the output voltage of a bandpass filter, to whose input the time function $f(t)$ is applied, provides the slope of the magnitude of the short-time spectrum for the frequency, to which the filter is matched. The envelope is obtained by a rectifier and followed by a smoothing lowpass filter. The envelope of the impulse response of the filter becomes effective as the weighting function. Thus, in the spectrum appears not only the Fourier transforms of the time function, but, in addition, the weighting function, too. In the case of a Fourier transformation in accordance with equation (1) of a pure sinusoidal oscillation, a line would appear in the spectrum. In the short-time spectrum, formed with a bandpass filter matched to the frequency of the sinusoidal oscillation, the line "degenerates" to the "filter shape" in the spectrum. The Fourier transform of a rectangular signal is the gap function $\frac{\sin x}{x}$. In the case of the ideal bandpass filter

with rectangular filter shape, therefore, the weighting function $w(t)$ in the time domain is equal to the gap function. Correspondingly, in the mathematical Fourier transformation of a rectangular section of a sinusoidal signal, the gap function appears in place of the spectral line.

The greater the filter bandwidth Δf , the smaller is the

Je größer die Filterbreite Δf desto kleiner ist die Filtereinschwingzeit, die Impulsantwort des Filters klingt schneller gegen Null ab, der Ausschnitt aus dem Zeitsignal Δt , der zum Kurzzeitspektrum transformiert wird, ist kürzer. Es gilt

$$\Delta t \cdot \Delta f \approx 1 \quad (5)$$

3. Parallelfilter-Verfahren

Ein häufig verwendetes Verfahren zur Messung des Amplitudenspektrums einer Zeitfunktion ist charakterisiert durch die Verwendung einer Vielzahl parallel geschalteter Bandfilter. Das zu untersuchende Signal liegt gleichzeitig am Eingang aller Filter. Haben alle n Filter die gleiche Bandbreite Δf und schließen aneinander an, so wird der Frequenzbereich $n \cdot \Delta f$ überstrichen. Gebräuchlich ist jedoch, daß die Filterbreiten mit der Ordnungszahl zunehmen, und zwar ist

$$\frac{\Delta f_{n+1}}{\Delta f_n} = C. \quad (6)$$

Ist $C = 2$, handelt es sich um Oktavfilter-, bei $C = \sqrt{2} \approx 1,26$ um Terzfilter-Analysatoren. Abb. 4 zeigt das Blockschaltbild eines Terzfilter-Analysators in einer Bauform als »Echtzeit-Analysator«, die für Sprachuntersuchungen besonders geeignet ist. Das vom Mikrofon oder einem Signalspeicher kommende Sprachsignal wird direkt dem Analysator-Eingang zugeführt und fortlaufend in Kurzzeitspektren umgewandelt.

In Abb. 4 folgt den parallel geschalteten Terzfiltern je ein Detektor und ein Integrator, die der Ermittlung der Umhüllenden der Filterausgang-Spannung dienen. Die Halteschaltungen ermöglichen zusätzlich zum kontinuierlichen Betrieb, einzelne Spektren auf ein externes Kommando hin zu speichern oder auch die Maximalwerte in den einzelnen Kanälen aus einer Spektrenfolge festzuhalten. Ein elektronischer Multiplexer fragt die Kanäle zyklisch ab und bringt das Spektrum auf einem Bildschirm zur Darstellung. Zusätzlich können einzelne Pegelwerte digital angezeigt werden, außerdem wird der Gesamtpegel mit den verschiedenen Bewertungen A, B, C und D sowie linear ermittelt. Bei Wahl der Darstellungsart »Spektrum Halten« oder »Maximum Halten« läßt sich ein Spektrum ausschreiben (Abb. 5). Zu beachten ist, daß die Verwendung breiterer Filter bei höheren Frequenzen, was dem Frequenz-Auflösungsvermögen des menschlichen Ohres angepaßt ist, zur Folge hat, daß die Bewertungsfunktionen $w(t)$ in allen Kanälen unterschiedlich sind. Die Filter der Mittenfrequenzen 100 Hz und 10 kHz haben Filterbreiten von 23 bzw. von 2300 Hz, entsprechend unterscheiden sich die Einschwingzeiten und damit die Bewertungen des Zeitsignals um den Faktor 100. Einen Ausgleich schaffen die Integratoren, deren Zeitkonstanten stets länger als die Einschwingzeiten und für zumindest eine Gruppe von mehreren Filtern konstant sind. Bei dem im

response time of the filter; the pulse response of the filter decays more rapidly towards zero, the portion from the time signal Δt , which is transformed towards the short-time spectrum, is shorter.

The expression

$$\Delta t \cdot \Delta f \approx 1 \quad (5)$$

applies.

3. Parallel filter method

A frequently employed method for the measurement of the amplitude spectrum of a time function is characterized by the use of a considerable number of bandpass filters arranged in parallel.

The signal to be investigated is applied simultaneously to the input of all the filters. If all n filters have the same bandwidth Δf and are arranged in sequence, the frequency range $n \cdot \Delta f$ is covered. It is, however, usual that the filter bandwidths increase with the ordinal number,

$$\frac{\Delta f_{n+1}}{\Delta f_n} = C \quad (6)$$

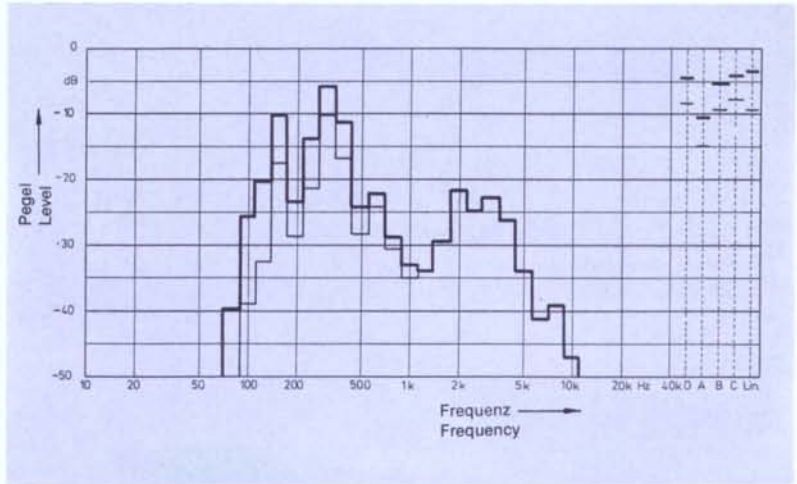
applying.

If $C = 2$, we are dealing with an octave filter, with $C = \sqrt{2} \approx 1,26$ we are dealing with 1/3 octave filter analyzers. Fig. 4 shows the block circuit diagram of a 1/3 octave filter analyzer designed as a "real-time analyzer", which is particularly suitable for investigations into speech. The speech signal coming from the microphone or from a signal store, is applied directly to the input of the analyzer and continuously converted into short-time spectra.

In Fig. 4 the 1/3 octave filters arranged in parallel, are each followed by a detector and an integrator which serve to establish the envelope of the filter output voltage. The hold circuits make possible, in addition to continuous operation, the storage of individual spectra in response to an external command, or the retention of the maximum values in the individual channels from a sequence of spectra. An electronic multiplexer scans the channels cyclically and represents the spectrum on a display screen. In addition, individual level values can be displayed digitally; moreover, the total level is established with the various weightings A, B, C and D and also linearly. If the mode of representation "Store" or "Max" is selected, a spectrum can be written out (Fig. 5). Account should be taken of the fact that the use of broader filters at higher frequencies - matched to the frequency resolution property of the human ear - results in variations in the weighting function $w(t)$ in all the channels. The filters of the centre frequencies 100 Hz and 10 kHz have filter bandwidths of 23 and 2300 Hz respectively, the response times are varying accordingly and thus also the weightings of the time

Abb. 5 Pegelerterspektren des Wortes »Ring« bei der Darstellungsart »Maximum Halten«, Zeitkonstantenwahl: »Sinus«, »Rauschen Schnell«

Fig. 5 Level 1/3 octave spectra of the word "Ring" with the mode of representation "maximum hold", time constant selection: "sinus", "noise, fast"



Beispiel der Abb. 5 verwendeten Gerät ließ sich die Zeitkonstante »Sinus«, »Rauschen schnell« und »Rauschen langsam« wählen. Im ersten Fall beträgt die Zeitkonstante zwischen 200 Hz und 40 kHz 0,2 s und steigt mit tiefer werdender Frequenz auf 2 s bei 20 Hz. Bei »Rauschen schnell« beträgt sie 0,2 s oberhalb 2 kHz und nimmt unterhalb kontinuierlich zu. In Abb. 5 wurden Spektren des Wortes »Ring« in diesen beiden Analysator-Einstellungen gewonnen. Es zeigen sich Pegelunterschiede von über 10 dB, da ja Mittelwerte über unterschiedlich lange Zeitabschnitte des Sprachsignals gebildet wurden. Zu beachten ist, daß dies bei den tieferen Frequenzen auch für die einzelnen Terzpegelwerte innerhalb eines Spektrums gilt. Für einsilbige Wörter bietet sich an, die Darstellungsart »Maximum Halten« zu wählen, um ein dem Lautheitseindruck zuzuordnendes Spektrum zu erhalten. Die experimentelle Untersuchung ergab jedoch, daß zum Beispiel bei der Analyse des Wortes »Spott« nach Abb. 1 der maximale Terzpegelwert im Kanal mit der Mittenfrequenz 4 kHz bereits 0,1 s nach dem in Abb. 1 eingetragenen Nullpunkt der Zeitachse gespeichert wurde, bei 1 kHz nach 0,35 s und bei 8 kHz nach 0,6 s.

Diese Verhältnisse müssen bei der Interpretation der Spektren sorgfältig berücksichtigt werden. Ein Pegelerterspektrum, das durch Integration über eine wählbare und in allen Kanälen gleiche Zeit, zum Beispiel also über ein Wort entsteht, läßt sich nur mit Analysatoren erreichen, die einen eigenen digitalen Mittelwertbildner besitzen, oder mit den neuesten Entwicklungen, den digitalen Terzanalysatoren.

Zur Erfassung der Lautheit von Sprachsignalen wurde ein spezieller Analysator entwickelt, der im Prinzip auf den Terz/Oktavanalysatoren fußt, jedoch zusätzlich Verdeckungseffekte berücksichtigt. Abb. 6 zeigt qualitativ ein

signal by the factor 100. A compensation is accorded by the integrators, the time constants of which are always longer than the response times and, at least for a group of several filters, constant. In the equipment employed in the example illustrated in Fig. 5, the time constant "Sine", "Fast Random" and "Slow Random" could be selected. In the former case, the time constant is 0.2 s between 200 Hz and 40 kHz and increases with lower frequencies to 2 s at 20 Hz. In "Fast Random" it is 0.2 s above 2 kHz and below increase continuously. In Fig. 5, spectra of the word "Ring" were obtained at these two settings of the analyser. Level differences of more than 10 dB appeared, since mean values were obtained over varyingly long time sections of the speech signal. It should be taken into account that this applies at the lower frequencies, also to the individual 1/3 octave levels within a spectrum. For monosyllabic words, it would appear expedient to select the mode of representation "Max" in order to obtain a spectrum that is to be oriented to the impression of loudness. The experimental investigation showed, however, that, for example in the analysis of the word "Spott" according to Fig. 1, the maximum 1/3 octave level in the channel was stored at the medium frequency of 4 kHz already at 0.1 s after the zero point of the time axis in Fig. 1, at 1 kHz after 0.35 s and at 8 kHz after 0.6 s.

These findings must be carefully taken into consideration when interpreting the spectra. A 1/3 octave spectrum that is obtained by integration over a selectable time identical in all channels, for example, over a word, can be attained only with the aid of analyzers which possess a digital mean value detector, or with the aid of the most modern developments, the digital 1/3 octave analyzers.

For the determination of the loudness of speech signals, a special analyzer was developed which, in principle, bases on

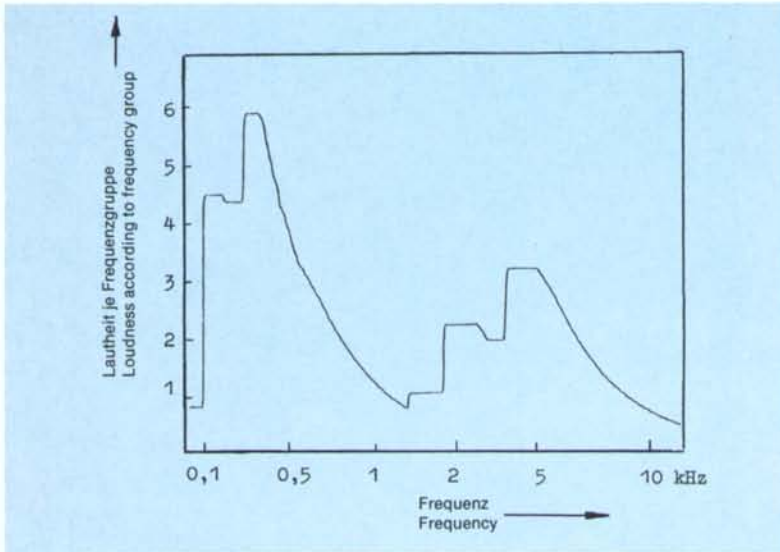


Abb. 6 Lautheitsdichtespektrum des Wortes »Ring«, Darstellungsart »Maximum Halten«

Fig. 6 Loudness density spectrum of the word "Ring" mode of representation "maximum hold"

Lautheitsdichte-Spektrum des Wortes »Ring«. Im unteren Frequenzbereich werden Oktav-, im oberen Terzfilter verwendet. Auch hier findet sich die Möglichkeit als Darstellungsart »Maximum Halten« zu wählen. Es wird hier jedoch das Spektrum gespeichert, das zum maximalen Gesamtlautheitswert zugehörig ist. Abb. 5 läßt sich daher durch die Unterschiede im Zeitbereich nicht in ein Abb. 6 entsprechendes Lautheitsdichte-Spektrum umrechnen.

(Fortsetzung folgt)

the 1/3 octave/octave analyzers, but in addition takes into account masking effects. Fig. 6 shows, qualitatively, a loudness density spectrum of the word "Ring". In the lower frequency range, octave filters were used and 1/3 octave filters in the upper range. Here, too, it is possible to select "Store Max" as a mode of representation. Here, however, the spectrum stored is that associated with the maximum total loudness value. Fig. 5, therefore, cannot be converted into a loudness density spectrum corresponding to Fig. 6, on account of the differences in the time range.

(To be continued)

Literatur/References

- [1] Brinkmann, K.: Die Neuaufnahme der »Wörter für Gehörprüfung mit Sprache«. Zts. f. Hör.-Ak. (J. audiol. Techn.) 13 (1974), 1.
- [2] Tscheschner, W.: Untersuchungs-
ergebnisse der Sprachanalyse. Zts. f. Phonetik 19 (1966), 141.
- [3] Flanagan, J. L.: Speech analysis synthesis and perception. Springer-Verlag, Berlin, Heidelberg, New York 1965.
- [4] Goydke, H.: Untersuchungen über die Spektral- und Autokorrelationsanalyse von Sprache zum Zweck der Sprecheridentifizierung. Dissertation TU Braunschweig 1974.